УДК [illegible]

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ОСНОВНЫХ КОМПОНЕНТ ИНФОРМАЦИОННО-ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ СИСТЕМЫ УПРАВЛЕНИЯ ЗНАНИЯМИ

А.Ф. Тузовский, И.А. Васильев, М.В. Усов

Институт "Кибернетический центр" ТПУ
E-mail: TuzovskyAF@kms.cctpu.edu.ru

Рассмотрен подход к реализации информационно-программного обеспечения системы управления знаниями с использованием языка представления знаний OWL и системы логического вывода RACER, использующей дескриптивную логику в качестве формализма представления знаний.

Дисциплина управления знаниями (УЗ) основывается на понятии "знание". В этом отношении УЗ отличается от управления информацией и не сводится к системам управления информацией (Content Management System) или поисковым системам. Знание – это такой продукт деятельности интеллекта человека, который получается в результате познания человеком закономерностей окружающего мира и их систематизации. Получая и применяя знания, человек может приспосабливаться к постоянно меняющимся условиям окружающей его среды. Считается, что знания неотделимы от человека. То есть они существуют как некоторый феномен в его голове. При передаче знаний от человека (источника) к человеку (приемнику) они проходят через цепочку преобразований "знание – информация – знание". Информация отличается от знаний в том смысле, что она не содержит всего контекста, в котором она существует в голове человека. Знания целостны, а информация – отрывки знаний. Знания могут быть успешно переданы от источника к приемнику посредством информации, если информация достаточно хорошо описывает контекст, в рамках которого существует передаваемое знание, и если приемник готов к интерпретации контекста и знаний. В терминах УЗ информацию, представленную в виде, удобном для хранения (текст, звук, видео), называют "явными знаниями". А знания в голове человека принято назвать "неявными знаниями", потому что они не выражены ни одним из способов, пригодных для их хранения и передачи.

Внутри компании знания как объект управления имеют свой жизненный цикл. Любая компания зарождается на определенном уровне знаний своих сотрудников, функционирует, используя и пополняя корпоративные знания (о продукте, о технологиях, о потребителях и рынке), являясь одновременно и получателем, и источником такого рода знаний. В компании и вне ее происходит неп-

рерывный процесс движения информации и знаний. Знания поступают в компанию в разнообразных формах (набор персонала, повышение квалификации, изучение специальной литературы, приобретение лицензий), формируются и развиваются в процессе проектной, инженерно-технической, производственной и маркетинговой деятельности, передаются вместе с товарами и услугами. Таким образом, в каждой компании, вне зависимости от сферы ее деятельности, вне зависимости от степени ее взаимодействия с внешним миром, осуществляется непрерывный процесс движения знаний большей или меньшей степени интенсивности, то есть осуществляется некий процесс производства, обобщения и распространения знаний, в том числе через продукцию и услуги потребителям.

В жизненном цикле знаний можно выделить три основных этапа: а) выявление потребности в знаниях; б) производство (создание) знаний; в) потребление (утилизация) знаний. В различных моделях, эти этапы включают разное количество процессов. С учетом того, что знания компании являются конкурентным преимуществом, необходима целенаправленная деятельность по стимулированию накопления и использования знаний в компании. **Управление знаниями в компании заключается в управлении жизненным циклом знаний за счет ведения целенаправленной деятельности по следующим направлениям (процессам в жизненном цикле знаний)**: идентификация существующих знаний; создание новых знаний; хранение знаний; распространение знаний; и стимулирование использования знаний.

Управление знаниями должно быть подчинено следующим принципам [1]:

– Организация вокруг бизнес-процессов. Основное усилие должно быть направлено на выявление возможностей для усовершенствования существующих в компании бизнес-процессов. Процессы по управлению знаниями должны "переплетаться" с существующими бизнес-процессами, способствуя накоплению и распространению знаний в компании. Так как бизнес-процессы реализуются сотрудниками компании, то управление знаниями должно учитывать специфические потребности в информации каждого из сотрудников. Таким образом, задачи по управлению знаниями компании формируются из задач бизнес процессов и информационных потребностей реализующих их сотрудников компании.
– Сохранение знаний и поддержка общения. В компании с развитой организационной структурой большое значение приобретает процесс распространения информации о результатах работы подразделений и отделов компании. Посредством управления знаниями нужно достичь определенного уровня осведомленности сотрудников о ходе бизнес-процессов, о результатах коллег. Наибольшая продуктивность сотрудника достигается при условии, что он может выполнять свои обязанности достаточно автономно, независимо, но при этом он осведомлен об общем состоянии бизнес-процесса в котором он участвует. Этого можно добиться за счет четкого распределения обязанностей (ролей) сотрудников во всех бизнес-процессах и создания условий для распространения релевантной информации.
– Ориентация на будущее. Стратегический менеджмент задает основной курс для развития компании, выявляет ее сильные и слабые стороны, анализируя успехи и неудачи в достижении поставленных целей. Долгосрочное планирование ориентировано на будущее и является неотъемлемой частью процесса управления. Оно должно осуществляться с учетом существующих в компании бизнес-процессов, а доведение стратегии до сотрудников компании требует отдельного внимания. Задача управления знаниями заключается в согласовании желания служащих сохранить автономность с необходимостью корректировки бизнес-процессов.
– Поддержка целей и задач компании. Компания в качестве одного из способов достижения поставленных задач использует формирование образа компании и корпоративной культуры. Корпоративные регламенты специфицируют бизнес-процессы, заставляя их работать на достижении поставленных целей. Значимость корпоративной культуры можно повысить, если оценивать вклад каждого сотрудника. Для этого необходима система метрик и показателей, которые указывали бы на степень достижения подразделениями и отдельными сотрудниками поставленных целей и задач, что, в конечном счете, влияет на достижение целей и задач компании.
– Продвижение инноваций. Отдельные сотрудники, коллективы людей и целые компании часто сталкиваются с необходимостью адаптироваться к меняющейся окружающей среде. Для этого приходится затрачивать усилия на формирование новых подходов к функционированию, которые, возможно уже были придуманы и опробованы. Управление знаниями должно быть сосредоточено на совершенствовании распространения инноваций, способствующих повышению эффективности отдельных сотрудников и целых бизнес-процессов.

Итак, **управление знанием имеет место тогда, когда могут быть созданы условия для умножения и распространения знаний и когда эти условия целенаправленно создаются**. На практике подходы к управлению знаниями могут различаться. Их можно сгруппировать следующим образом: а) организационный; б) технологический подход.

Сторонники организационного подхода считают, что благоприятная среда для создания и распространения знаний в компании может быть создана только организационными мерами. Только прямое общение людей способно приумножить знание. Информация, передаваемая посредством, например, документов, содержит лишь часть контекста, кото-

рый должен быть дополнен общением, чтобы передача знаний прошла успешно. Общаясь, человек может всегда уточнить что-то у собеседника, спросить значение термина, если у него возникают вопросы или сомнения. Таким образом, передаваемая информация всегда "обрамлена" в нужный контекст, и поэтому способна нести в себе знания. В рамках данного подхода компьютерная техника и программные системы имеют лишь второстепенное значение.

Технологический подход в один ряд с организационными мерами ставит применение вычислительной техники и информационных технологий. Хотя информация, например, в виде документа еще не гарантирует передачу заложенных в ней знаний, она имеет ряд преимуществ перед знаниями. Информация, представленная в электронном виде, может сколько угодно долго храниться, ее удобно копировать и передавать. Это говорит о том, что нужно стремиться активно использовать информационные технологии в управлении знаниями. А возможность их использования на сегодняшний день тесно связана с развитием онтологического подхода к представлению знаний. Онтология представляет знания человека в виде набора понятий об окружающем мире и отношений между этими понятиями. Понятия и отношения образуют сеть (граф). Онтология служит промежуточным звеном между ЭВМ, которая содержит знания человека в виде информации, и человеком, который получает эту информацию и "извлекает" из нее знания. Онтология в некотором смысле заменяет автора информации, обрамляя ее в контекст, необходимый для усвоения информации другим человеком. То есть онтологию можно назвать "картой знаний", которая разъясняет смысл отдельных информационных блоков, обрамляет их в целостный контекст.

Мы поддерживаем технологический подход к управлению знаниями. По определению, система управления знаниями (СУЗ) – это набор средств (подходов, организационных мероприятий, техническое и программное обеспечение) для систематической, целенаправленной работы по обеспечению сотрудников организации всеми видами полезных знаний, которые имеются как в самой организации, так и за ее пределами. Базой для реализации СУЗ должны стать информационные технологии. Информационные технологии способны обеспечить скорость обработки и широту охвата проблемной области, а организационные меры – эффективное функционирование программно-технического комплекса.

В качестве базовой технологии поддержки управления знаниями компании выступает информационный портал предприятия (ИПП). Как определение того, что является информационным порталом предприятия, так и функциональность программных продуктов этого класса варьируются. Например, одно из первых определений, данных в 1998 г., звучит следующим образом:

Информационные порталы предприятия является приложениями, которые помогают компаниям "высвободить" информацию, хранящуюся как внутри компании, так и вне ее, предоставляют пользователям единую точку доступа к персонифицированной информации, необходимой для принятия решений... Это слияние различных программных систем, обеспечивающих консолидацию, управление, анализ и распространение информации как внутри компании, так и за ее пределами.

Существуют различные классификации ИПП на основании множества критериев, в том числе и на основании функциональности. Можно привести обобщенный список их функциональных возможностей:

- интеграция данных и приложений в организации;
- категоризация данных по различным критериям;
- организация коллективной работы сотрудников организации;
- персонификация пользовательского доступа;
- публикация документации с применением процесса одобрения;
- предоставление сотрудникам различных средств взаимодействия друг с другом;
- поиск по всевозможным источникам информации;
- широкий набор административных функций.

Решение о том, когда следует внедрять ИПП, принимается отдельно каждой компанией. Существуют различные методики оценки "возврата инвестиций" от внедрения ИПП. Практика показывает, что вышеупомянутые функциональные возможности действительно **способны повысить производительность сотрудников компании и снизить стоимость информации**. Как отмечают аналитики Gartner Group [2], сектор рынка ИПП растет, в нем выделяются поставщики-лидеры, клиентами которых являются крупные и успешные компании по всему миру. Так, по состоянию на март 2004 г. среди лидеров на рынке ИПП значились IBM, SAP, Orade, BEA Systems, Plumtree Software, Sun Microsystems, PeopleSoft, Vignette. Например, из крупных клиентов компании IBM можно отметить концерн Volkswagen, корейскую вещательную корпорацию KOBACO, немецкую фармацевтическую корпорацию Schering AG.

Специфика ИПП такова, что, во-первых, составляющие его компоненты – отдельные сервисы и целые приложения – тесно связаны с бизнес-процессами компании, и поддерживается интеграция различных источников данных, что позволяет объединить разрозненные хранилища в единое информационное пространство компании. Во-вторых, ИПП удобно разворачивать и поддерживать. В-третьих, доступ к ИПП возможен как изнутри корпоративной сети компании, так и из-за ее пределов, что удобно для сотрудников компании, ее клиентов и партнеров.

Информационный портал предприятия, который является частью системы управления знаниями компании, называют порталом знаний предприятия (ПЗП) [3]. В рамках СУЗ основными зада-

чами ПЗП являются хранение, структуризация и распространение знаний. Эти задачи решаются с помощью **онтологий** и **семантических метаданных**.

Как уже было сказано, онтология представляет собой сеть, образованную понятиями и связывающими их отношениями. Например, "человек" и "возраст" – это понятия, а "иметь возраст" – это отношение. Они образуют цепочку "человек – иметь возраст – возраст". Понятия являются узлами сети, а отношения – дугами, связывающими понятия. Множество понятий и отношений образуют сеть. Такая сеть является способом представления знаний человека. Знания, взятые **в полном объеме**, у каждого человека уникальны. Однако по какому-то конкретному вопросу группа людей может прийти к согласию, то есть разделять одинаковые взгляды и иметь одинаковые знания. Поэтому можно спроектировать такую онтологию, которая была бы проблемно-ориентированная и представляла знания, которыми обладает не отдельный человек, а с которыми "согласна" группа людей [4]. Именно такие онтологии и используются в ПЗП. Онтология в ПЗП выполняет роль осей, описывающих пространство, а семантические метаданные, прикрепленные к каждому информационному объекту ПЗП, выполняют роль координат объекта в этом пространстве. В качестве информационных объектов чаще всего выступают документы, аудио- и видео-файлы, профили пользователей и, конечно, поисковые запросы.

Таким образом, если каждый объект имеет свои координаты или другими словами, свой семантический образ, то можно вычислять степень близости двух объектов или так называемую "семантическую схожесть", сравнивая семантические образы объектов. Например, сравнивая семантический образ профессиональных интересов сотрудника компании и семантический образ книги, можно определить, должен ли ПЗП предложить эту книгу для прочтения или нет. Семантический образ той же книги можно сравнить с семантическими образами рубрик хранилища документов, чтобы автоматически категорировать данную книгу. Для реализации таких возможностей необходимо **определить методику семантического сравнения** [5] и подходящую для нее **структуру метаданных**.

Методика семантического сравнения должна отражать логику рассуждений человека. Если мы хотим определить, насколько близко два понятия находятся друг от друга, то нужно понять, как это делает человек.

Взаимодействуя с окружающей средой, человек постоянно сталкивается с различными **экземплярами** объектов – людьми, явлениями природы и т.п. Находя в экземплярах общие черты, выражающиеся в значениях каких-либо свойств экземпляров и в отношениях с другими экземплярами, человек объединяет экземпляры в **классы**, формируя у себя в голове тем самым обобщающие **понятия**. Понятия абстрактны, характеризуются отношениями с другими понятиями. Экземпляры конкретны, характеризуются отношениями с другими экземплярами. Человек относит экземпляр к одному или более классу понятий, то есть экземпляр является представителем одного или более понятий. Человек может сравнивать между собой однотипные элементы – понятия с понятиями или экземпляры с экземплярами. Сравнение экземпляра с понятием нужно тогда, когда класс экземпляра неизвестен, и человек хочет понять, к какому понятию относится экземпляр. Понятия в голове человека упорядочиваются в иерархию – таксономию понятий, – в которой каждое понятие является подклассом одного или более других понятий. Предполагается, что исходя из таксономии, человек определяет схожесть понятий, а схожесть экземпляров – исходя из отношений с другими экземплярами.

Типы схожести:

1. **Схожесть между понятиями**. Как уже было сказано, схожесть двух понятий измеряется пересечением множеств, состоящих из отношений каждого из понятий с другими понятиями. Чем больше "одинаковых" отношений имеют оба понятия, тем больше они похожи. Анализируя таким способом схожесть понятий, можно упорядочить все понятия в таксономию – иерархию типа "класс-подкласс". Одно понятие ($П_1$) является подклассом другого понятия ($П_2$), если у $П_1$ есть такие же отношения с другими понятиями, как у $П_2$, и у $П_1$ есть еще дополнительные отношения. Есть и другие условия формирования иерархии класс-подкласс. Таксономия, таким образом, является производной от отношений между понятиями. Правильнее оценивать схожесть понятий по их отношениям с другими понятиями. Но часто бывает так, что одно понятие является подклассом другого, но неизвестно на основании каких отношений. В такой ситуации схожесть понятий не может быть оценена по отношениям. Таксономия же позволяет определить схожесть понятий в любой ситуации. Предлагается ввести две характеристики схожести между понятиями – **классовое расхождение** (P_K) и **таксономическую схожесть** (C_T). Классовое расхождение важно при поиске ответов на поисковые запросы и является основным показателем. Таксономическая схожесть предоставляет дополнительные возможности сравнения объектов в ПЗП.

   Значение P_K предлагается измерять в диапазоне [–1;1]. $P_K(П_1,П_2)=0$, если понятия $П_1$ и $П_2$ эквивалентны. $P_K(П_1,П_2)>0$, если понятие $П_1$ является более общим понятием, чем $П_2$. $P_K(П_1,П_2)<0$, если понятие $П_1$ является менее общим понятием, чем $П_2$. Таким образом, классовое расхождение несимметрично, $P_K(П_1,П_2)=–P_K(П_2,П_1)$.

   Значение C_T предлагается измерять в диапазоне [1;0). $P_K(П_1,П_2)=1$, если понятия $П_1$ и $П_2$ эквивалентны. $P_K(П_1,П_2)$ всегда больше 0. Вычисление основывается на нахождении наименьшего общего суперкласса ($П_{НО}$) для понятий $П_1$ и $П_2$.

$$C_T = \frac{2 - P_K(П_{НО}, П_1) - P_K(П_{НО}, П_2)}{2}.$$

Таксономическая схожесть симметрична, $P_K(П_1,П_2)=P_K(П_2,П_1)$.

2. **Схожесть между экземплярами понятий**. Измерять схожесть экземпляров $Э_1$ и $Э_2$ предлагается на основании проверки эквивалентности значений отношений, которые заданы как для $Э_1$, так и для $Э_2$. Для этого вводится **показатель реляционной схожести** СР($Э_1$,$Э_2$) диапазон значений которого равен [0;1]. СР($Э_1$,$Э_2$)=0, если значения отношений экземпляров не пересекаются, и СР($Э_1$,$Э_2$)=1, если значения отношений экземпляров пересекаются полностью. Реляционная схожесть симметрична, СР($Э_1$,$Э_2$)=СР($Э_1$,$Э_2$).

   Измерение схожести между экземплярами имеет смысл, если понятия этих экземпляров находятся в зависимости класс-подкласс. Исходя из этого, показатель схожести экземпляров вычисляется следующим образом:

$$C_Э(Э_1, Э_2) = \frac{1 - P_K + C_P}{2}.$$

Из предложенной методики семантического сравнения следует **логическая структура метаданных**, описывающих объекты в ПЗП. Она должна включать в себя множество понятий, ассоциированных с объектом, а также множество триплетов "экземпляр – отношение – экземпляр".

Использование в ИПП таких семантических технологий как онтологии и семантические метаданные позволяет перейти от синтаксиса к семантике в обработке информации. Такой переход направлен на преодоление проблем, связанных со структуризацией и поиском информации [3]. ИПП, которые используют семантические технологии, построены в соответствии с принципами УЗ и решают задачи УЗ, являются ПЗП.

Для построения ПЗП на основе ИПП необходимы:

– методология создания и поддержки онтологий;
– язык описания онтологий;
– программные средства для работы с этим языком;
– язык описания семантических метаданных;
– методы вычисления семантической схожести объектов.

Существует множество языков описания онтологий и методологий проектирования онтологий. Для поддержки языков разработаны разнообразные программные средства – редакторы, хранилища, языки запросов и т.п. Кроме этого, от выбора языка зависит структура семантических метаданных и методы вычисления семантической схожести. Все это приводит к тому, что реализации порталов знаний могут очень сильно различаться [3].

В качестве языка описания онтологий мы используем **дескриптивную логику** класса "ALCQHIr+". Дескриптивные Логики (ДЛ) являются семейством языков представления знаний, которые определяют формальную модель для описания знаний предметной области и определяют синтаксис языка для записи моделей [6]. Все языки из семейства ДЛ являются подмножеством логики первого порядка (ЛПП) и различаются выразительной способностью и возможностями логического вывода. В отличие от ЛПП, все языки семейства ДЛ являются разрешимыми за конечное время. Так как гарантия ответа в ограниченное (конечное) время не предполагает, что ответ будет дан через приемлемый интервал времени, то исследование вычислительной сложности в конкретной ДЛ с разрешимыми задачами логического вывода, является важной проблемой. Разрешимость и сложность задач вывода зависит от выразительной силы исследуемой ДЛ. С одной стороны, очень экспрессивные ДЛ наиболее вероятно имеют задачи вывода высокой сложности, требующие значительного времени. С другой стороны, очень слабые ДЛ (с эффективными процедурами вывода) могут не обладать достаточной выразительностью для представления важных понятий конкретной прикладной задачи. Исследование этого баланса между выразительностью ДЛ и сложностью их задач вывода было одной из наиболее важных проблем в исследованиях ДЛ. Все языки семейства ДЛ упорядочены по возрастанию выразительной способности и именуются специальными аббревиатурами, начиная с "AL" с добавлением других букв-кодификаторов (например, "ALCQ^{-1}r+").

ДЛ описывают знания предметной области, вначале, описывая необходимые понятия предметной области (его терминологию), а затем используя эти понятия для точного описания свойств объектов и индивидуумов (экземпляров), встречающихся в этой предметной области. Ключевой особенностью ДЛ, что отличает их от предшественников, таких как семантические сети и фреймы, является то, что это логики, т.е. формальные языки с хорошо определенной семантикой (смыслом). Другой их отличительной особенностью является упор на логический вывод, как основной способ использования: логический вывод позволяет выводить неявно представленные знания из знаний, которые явно содержатся в базе знаний. ДЛ поддерживают шаблоны логического вывода, которые встречаются во многих вариантах интеллектуальных систем обработки информации, и которые также используются людьми, для структурирования и понимания мира: классификация понятий и их представителей (экземпляров).

Основные идеи ДЛ:

– Синтаксическими строительными блоками являются: атомарные понятия (унарные предикаты), атомарные роли (бинарные предикаты), и представители (экземпляры, константы). Роли (свойства, отношения) являются самостоятельными элементами, которые затем могут быть связаны с понятиями.
– Понятия и роли можно объединяться в выражения для описания более сложных понятий с помощью конструкторов (операций).

– Между понятиями и между ролями можно задать отношения (какие понятия/роли являются тождественными, и какие понятия/роли включают другие понятия/роли).
– Выразительная мощь языка ограничивается тем, что он использует достаточно малый набор (достаточных для описания знаний) конструкторов для построения сложных понятий и ролей.
– Неявные знания о понятиях и экземплярах могут быть выведены автоматически с помощью процедур вывода. В частности, отношения включения (родовидовые отношения) между понятиями и ролями.

Дескриптивная логика класса "ALCQHIr+" для описания онтологии предметной области предоставляет следующие примитивы: атомарные понятия и роли, отрицание произвольных понятий, иерархия ролей, инвертированные роли, транзитивные роли, ограничение на значение роли вида $\leq n$R.C.

Мы считаем, что применение ДЛ **для целей управления** знаниями оправдано потому, что:
– ДЛ класса "ALCQHIr+" обладает высокой выразительной способностью и разрешима;
– ДЛ позволяет проверять таксономии понятий на непротиворечивость и автоматически строить их;
– отношение "класс-подкласс" является основополагающим в задачах поиска и классификации текстовой информации;
– существует достаточное множество прикладных программ для работы с ДЛ (языки описания, редакторы, системы логического вывода);
– существуют методологии проектирования онтологии, соответствующие особенностям ДЛ.

В связи с тем, что ПЗП является развитием ИПП, мы предлагаем создавать ПЗП как конгломерат ИПП и "пространства знаний" (рис. 1). Существующие в компаниях ИПП при необходимости могут быть дополнены семантическим ядром, которое дополняет стандартные функции ИПП – сбор, обработку, хранение, извлечение, преобразование и представление информации – семантическими сервисами, основанными на онтологии, семантических метаданных и методах оценки семантической близости.

Ниже приведена предлагаемая архитектура семантического ядра ПЗП (рис. 2). Ее основные элементы:
– хранилище онтологий (сервер онтологий);
– система логического вывода;
– объектная модель онтологии;
– хранилище семантических метаданных (сервер семантических метаданных);
– модуль семантического аннотирования;
– модуль семантического сравнения.

Сервер онтологий хранит используемые порталом онтологии, описанные на языке OWL, в реля-

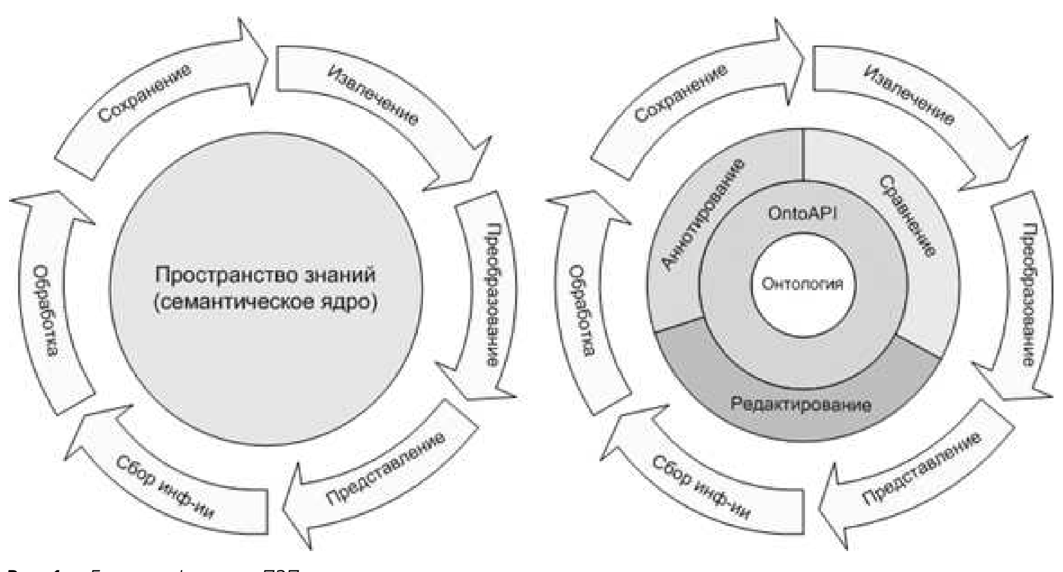


Рис. 1. *Базовые функции ПЗП*

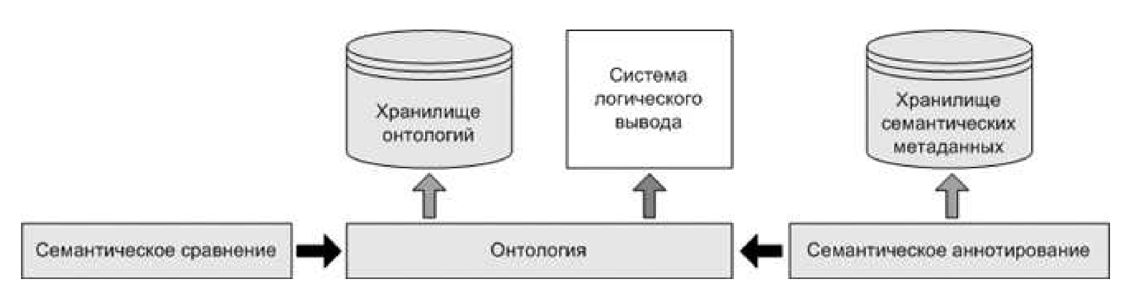


Рис. 2. *Схема семантического ядра*

ционной базе данных. Сервер онтологий предоставляет доступ к отдельным онтологиям через web-сервис, сохраняя и извлекая онтологию из хранилища целиком или по частям. Частью реализации сервера является конвертер из формата OWL в реляционную схему и наоборот. Язык OWL выбран потому, что он позволяет описать все примитивы дескриптивной логики, и он становится стандартом записи онтологий.

С сервером онтологий напрямую работает **объектная модель онтологии**, которая представляет понятия и отношения из онтологии в удобном для работы, объектно-ориентированном виде. Объектная модель онтологии обращается к серверу онтологий по протоколу SOAP (web-сервис), чтобы получить описание онтологии на языке OWL и представить его в виде объектов, таких как "понятие", "отношение", "свойство" и т.п. Чаще всего (хотя и не всегда), онтология, полученная от сервера онтологий, является верифицированной. То есть понятия в онтологии упорядочены в таксономию, и выполнять логический вывод не требуется. Однако дескриптивная логика позволяет работать с такими описаниями понятий, которые явно в онтологии не указаны. Для этого объектная модель онтологии взаимодействует с **системой логического вывода**, которая проверяет новые понятия на непротиворечивость и определяет их место в таксономии.

В качестве системы логического вывода для дескриптивной логики нами используется программный модуль RACER, созданный специалистами Гамбургского технологического университета. Этот модуль предоставляет возможность логического вывода для дескриптивной логики класса "ALCQHIr+". Это значит, что модуль поддерживает иерархию понятий и ролей, свойства, атрибуты, инверсию ролей, ограничения за значение домена и на количество отношений и т.д. Помимо этого поддерживается проверка непротиворечивости онтологии, а также логический вывод на экземплярах понятий (так называемый ABox) и одновременная обработка нескольких онтологий. Взаимодействие с системой логического вывода осуществляется по протоколу DIG (Description logic Interface Group).

Модуль семантического аннотирования предоставляет возможность ручного и полуавтоматического связывания понятий из онтологии предметной области с текстовой информацией. Процесс связывания называется аннотированием и основывается на морфологическом и синтаксическом анализе текста и на поиске понятий в онтологии. Для записи метаданных используется язык RDF(S). Полученные аннотации (семантические образы) сохраняются в **хранилище семантических метаданных**, обращение к которому осуществляется по протоколу SOAP (web-сервис).

Модуль семантического сравнения использует описанную выше методику вычисления семантической схожести и зависит от системы логического вывода, которая строит таксономию понятий и проверяет целостность базы знаний.

СПИСОК ЛИТЕРАТУРЫ

1. Collins H. Enterprise knowledge portals: next generation portal solutions for dynamic information access, better decision making and maximum results. — N.Y.: AMACOM, 2003. — 403 p.
2. Phifer G., Valdes R., Gootzit D., Underwood K., Correia J., Andrews W. Magic Quadrant for Horizontal Portal Products, 2004. Gartner Group. — 10 p. http://mediaproducts.gartner.com/reprints/ibm/120327.html
3. Lausen H., Stollberg M., Hernández R.L., Ding Y., Sung-Kook H., Fensel D. Semantic Web Portals — State of the Art Survey, 2004. — 47 p. http://www.deri.at/publications/techpapers/documents/DERI-TR-2004-04-03.pdf
4. Borst W.N. Construction of engineering ontologies for knowledge sharing and reuse. Ph.D. thesis. 1997. — 243 p.
5. Stojanovic N., Mädche A., Staab S., Studer R., Sure Y. SEAL — A framework for developing semantic portals. // Proc. of the I Int. ACM Conf. on knowledge capture. — Victoria, 2001. — 8 p.
6. The Description Logic handbook: theory, implementation, applications / Ed. F. Baader, Cambridge University Press, 2003. — 564 p.